



HAL
open science

An overview of a posteriori error estimation and post-processing methods for nonlinear eigenvalue problems

Geneviève Dusson, Yvon Maday

► **To cite this version:**

Geneviève Dusson, Yvon Maday. An overview of a posteriori error estimation and post-processing methods for nonlinear eigenvalue problems. 2023. hal-04010911v2

HAL Id: hal-04010911

<https://hal.science/hal-04010911v2>

Preprint submitted on 3 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An overview of *a posteriori* error estimation and post-processing methods for nonlinear eigenvalue problems

Geneviève Dusson^{*1} and Yvon Maday^{†2}

¹Université de Franche-Comté, Laboratoire de mathématiques de Besançon,
UMR CNRS 6623, 16, route de Gray, 25000 Besançon, France

²Sorbonne Université, CNRS, Université Paris Cité, Laboratoire Jacques-Louis Lions
(LJLL), F-75005 Paris, France et Institut Universitaire de France.

Abstract

In this article, we present an overview of different *a posteriori* error analysis and post-processing methods proposed in the context of nonlinear eigenvalue problems, e.g. arising in electronic structure calculations for the calculation of the ground state and compare them. We provide two equivalent error reconstructions based either on a second-order Taylor expansion of the minimized energy, or a first-order expansion of the nonlinear eigenvalue equation. We then show how several *a posteriori* error estimations as well as post-processing methods can be formulated as specific applications of the derived reconstructed errors, and we compare their range of applicability as well as numerical cost and precision.

1 Introduction

Nonlinear eigenvalue problems occur in many mathematical models used in science and engineering such as the calculation of the vibration modes of a mechanical structure in the framework of nonlinear elasticity, the ground state of the Gross–Pitaevskii equation describing the steady states of Bose–Einstein condensates [50], or of the Hartree–Fock and Kohn–Sham equations ([43] used to calculate ground state electronic structures of molecular systems in quantum chemistry and materials science (see [9] for a mathematical introduction).

The approximation of the solutions to such problems is of major importance and different methods to compute them are proposed depending on the various applications at stake. These approximation methods are based on different ingredients. The first one is related to the notion of degrees of freedom, associated with the basis sets used to approximate the solutions of these problems, which leads to discrete problems that can be solved, eventually, on a computer. Second, the resulting discrete problems are, per force, nonlinear; hence efficient algorithms must be designed to solve these problems accurately with a reasonable computational cost. Regarding the problems studied in this article, two main families of methods exist. First, there exist algorithms

^{*}genevieve.dusson@math.cnrs.fr

[†]yvon.maday@sorbonne-universite.fr

directly minimizing the energy functional, such as saddle point problems solved with Newton-type methods [2, 3], projected Sobolev gradient flow methods [39, 27, 32, 34, 52, 62, 1]. The second type of methods solve the corresponding first-order Euler–Lagrange equations, in the form of a nonlinear eigenvalue problem, based on the Self-Consistent Field (SCF) algorithm [47, 53, 5, 15, 22, 55]. In [14], an analysis comparing direct minimization methods and SCF algorithms is proposed, see also the references therein.

These two approximation ingredients must be well tuned so that the approximate solutions are close enough to their corresponding exact ones. Moreover the exact solutions being unknown, arguments to estimate the error between exact and approximate solutions need to be proposed. Such arguments often rely on so-called *a priori* and *a posteriori* analysis, where the *a priori* analysis aims at providing convergence rates of the approximate solutions towards the exact ones, while the *a posteriori* analysis derives error bounds that should only depend on quantities computable from the approximate solutions the accuracy of which one wants to qualify. These last years, a large number of articles about the numerical analysis of approximation of solutions to problems set in the form of a partial differential equation in fields like fluid mechanics and solid mechanics have been published, see e.g. the monograph [56] on this subject, focusing mainly on characterizing the number of degrees of freedom necessary to reach a given accuracy.

In the context of the current article, the *a priori* analysis for the class of nonlinear eigenvalue problems is quite recent and relies on the papers [9, 64, 7, 18, 40, 8, 17] and the references therein. These articles consider the discretization error, that is the error due to the use of a given number of degrees of freedom in the approximation of the problem of interest. They provide convergence rates i.e. upper bounds for the decay rate of the error – both on the eigenvectors and the eigenvalues – when the number of degrees of freedom increases.

These *a priori* approaches allow first to state that: i) it is possible to achieve a satisfying approximation, provided that computing resources are sufficient; this is the notion of convergence, then, ii) for a problem in dimension d , multiplying the number of degrees of freedom by 2^d leads to an error decay by a factor 2^r where r is related to the order of convergence of the method. Such results are nevertheless insufficient in general since most of the time, the affordable discretizations are limited due to the high computational cost of the methods which may be polynomial in the number of degrees of freedom e.g. $[2^d]^p$, where p can be $\simeq 5$ or 10 . Hence the interest rather goes in estimating, once a computation has been done with a given discretization, the magnitude of the error. This is precisely the aim for which *a posteriori* approaches (estimators and indicators) have been designed. These analyses do not usually give rates of convergence but instead provide actual figures that are upper bounds (and also, most of the time now, lower bounds) of the errors between exact and approximate quantities of interest such as in this context: energies, eigenvalues, wavefunctions (eigenfunctions). Also they may allow for an optimization of the choice of the degrees of freedom (such as basis functions) in order to minimize the computational cost to reach a given accuracy.

As far as we know, the first article in the direction of *a posteriori* estimates for nonlinear eigenvalue problems is [46], where the *a posteriori* analysis of the Hartree–Fock problem was performed and upper and lower error bounds were proposed for the ground state energy. Other contributions have been proposed since, e.g. [21, 18, 20], where error bounds for nonlinear eigenvalue problems are presented in the context of finite element discretizations, see also *a posteriori* results for Fourier (planewave) discretizations in [11, 25].

Let us also mention two-grid methods which have been designed not directly to provide error bounds, but to obtain accurate solutions at a low computational cost. In these methods, a first step consists in computing a coarse solution by performing a full calculation – in this case the resolution of a nonlinear eigenvalue problem – with a limited number of degrees of freedom, which should be reasonably cheap because of the small number of involved degrees of freedom. The second step is to perform a simpler computation, e.g. to solve a linear eigenvalue problem or a boundary value problem, with a larger number of degrees of freedom, to improve the coarse solution. For nonlinear eigenvalue problems, two-grid methods have been proposed for example in [33, 6] for a Gross–Pitaevskii type equation and [12] for Kohn–Sham models, the latter relying on a perturbation method. We will see in the following that the analysis techniques used in these two-grid methods can be very similar to the ones used in the former *a posteriori* methods.

The different references quoted above have some methodology in common but are applied to different problems, possibly with different boundary conditions, and discretized with different methods. This makes a straightforward comparison of these approaches quite difficult. In this article, we intend to compare them in a unified framework, clarifying their main similarities and differences (see Table 1). To do so, we will present all the methodologies in the framework of a simple nonlinear eigenvalue problem. More precisely, we place ourselves in the periodic setting where thus the domain Ω is the unit cell of a periodic lattice \mathcal{R} of \mathbb{R}^d and $X = H_{\#}^1(\Omega)$, with $d = 1, 2$ or 3 . Then for all $s \in \mathbb{R}^+$ and $k \in \mathbb{N}^1$,

$$\begin{aligned} H_{\#}^s(\Omega) &= \left\{ v|_{\Omega}, v \in H_{\text{loc}}^s(\mathbb{R}^d) \mid v \text{ is } \mathcal{R}\text{-periodic} \right\}, \\ H_{\#}^{-s}(\Omega) &= [H_{\#}^s(\Omega)]' \quad (\text{dual space of } H_{\#}^s(\Omega)), \\ C_{\#}^k(\Omega) &= \left\{ v|_{\Omega}, v \in C^k(\mathbb{R}^d) \mid v \text{ is } \mathcal{R}\text{-periodic} \right\}. \end{aligned}$$

We then focus on a particular class of (linear and) nonlinear eigenvalue problems arising in the study of variational models of the form

$$I = \inf \left\{ \mathcal{E}(v), v \in X, \int_{\Omega} v^2 = 1 \right\}, \quad (1.1)$$

associated to an energy functional \mathcal{E} of the form

$$\mathcal{E}(v) = \frac{1}{2}a(v, v) + \frac{\mu}{2} \int_{\Omega} G(v^2(x)) dx, \quad v \in X, \quad (1.2)$$

where $\mu = 1$ (nonlinear case) is the case of interest but we may also have $\mu = 0$ that corresponds to the linear case, and where

1. the term in the integral involving a given function G depends the square of solution (v^2) to mimic the electronic structure calculation problems we are mostly interested in such as the Hartree–Fock and Kohn–Sham problems. In these problems, the nonlinearity depends on the electronic density ρ , which corresponds to v^2 when the energy functional depends only on one function in X . In what follows, we denote by g the derivative of G .

¹Note that $L_{\#}^2(\Omega)$ coincides with $L^2(\Omega)$

2. the bilinear form a is defined by

$$a(u, v) = \int_{\Omega} (A \nabla u) \cdot \nabla v + \int_{\Omega} V u v.$$

3. we make the following assumptions on the elements appearing in the energy

- $A \in (L^{\infty}(\Omega))^{d \times d}$; $A(x)$ is symmetric for almost all $x \in \Omega$;
 $\exists \alpha > 0$ such that $\xi^T A(x) \xi \geq \alpha |\xi|^2$, $\forall \xi \in \mathbb{R}^d$ and almost all $x \in \Omega$;
- $V \in L^p(\Omega)$ for some $p > \max(1, d/2)$;
- $G \in C^1([0, +\infty), \mathbb{R}) \cap C^2((0, \infty), \mathbb{R})$, $g = G'$, $g(0) = 0$ and $g' > 0$ on $(0, +\infty)$;
 $\exists 0 \leq q < 2$, $\exists C \in \mathbb{R}_+$ such that $\forall t \geq 0$, $|g(t)| \leq C(1 + t^q)$;
- $g'(t)t$ is locally bounded on $[0, +\infty)$.

There is no loss of generality in assuming in (1.3) that $g(0) = 0$ since the minimizers of (1.1) are not modified if $G(t)$ is replaced with $G(t) + ct$, due to the normalization constraint imposed on the solutions of (1.1).

It is well known (see e.g. Lemma 2 in [7]) that under these assumptions, problem (1.1) has exactly two minimizers u and $-u$, one of them, say u , being positive on Ω . In all what follows, u will be the positive minimizer of (1.1). Let us introduce the Fock operator, defined, for any $u \in X$ by $\mathcal{A}_u = D_v \mathcal{E}(u)$, i.e.

$$\mathcal{A}_u = -\operatorname{div} A \nabla + V + \mu g(u^2). \quad (1.4)$$

We then denote by \mathcal{A} the operator corresponding to the linear case, i.e. for $\mu = 0$, so that

$$\mathcal{A} = -\operatorname{div} A \nabla + V. \quad (1.5)$$

Writing the Euler–Lagrange equation of problem (1.1), we obtain that the function u is solution to the nonlinear eigenvalue problem

$$\forall v \in X, \quad \langle \mathcal{A}_u u - \lambda u, v \rangle_{X', X} = 0, \quad (1.6)$$

where the eigenvalue $\lambda \in \mathbb{R}$ is the Lagrange multiplier associated with the constraint $\|u\|_{L^2} = 1$. Equation (1.6), complemented with the constraint $\|u\|_{L^2} = 1$, reads

$$\begin{cases} \mathcal{A}_u u = \lambda u, \\ \|u\|_{L^2} = 1. \end{cases} \quad (1.7)$$

Note that, for any $v \in X$, \mathcal{A}_v is a linear self-adjoint operator on $L^2(\Omega)$ with form domain X . It can then be inferred from (1.7) that $u \in X \cap C^0(\overline{\Omega})$, $u > 0$ in Ω , and λ is the lowest eigenvalue of problem (1.7), called the ground state eigenvalue of \mathcal{A}_u . Note however that there exist cases where the lowest eigenvalue does not correspond to the lowest energy, such as a Gross–Pitaevskii equation with a rotating magnetic field, see [1, Section 6.1]. Regarding quantum chemistry equations, such as the Kohn–Sham or Hartree–Fock equations, a rule called Aufbau principle states that the ground state indeed corresponds to the lowest eigenvalues of the nonlinear operator \mathcal{A}_u . An important

point is that λ is a *simple* eigenvalue of \mathcal{A}_u . These results are classical; their proofs are recalled in [7].

We now consider a family of finite-dimensional subspaces $(X_\delta)_{\delta>0}$ of X (conforming approximation $X_\delta \subset X$) such that

$$\forall v \in X, \quad \lim_{\delta \rightarrow 0} \min_{v_\delta \in X_\delta} \|v - v_\delta\|_X = 0. \quad (1.8)$$

An example is the Fourier discretization. Let assume that $\Omega = (0, 2\pi)^d$. For any $k \in \mathbb{Z}^d$, we define the planewave e_k by $e_k(x) = (2\pi)^{-d/2} e^{ik \cdot x}$. We then consider the discretization space indexed by a parameter M that grows when the space grows (that is corresponding to $1/\delta$ above)

$$X_M = \text{Span} \left\{ e_k, k \in \mathbb{Z}^d, |k| \leq M \right\}, \quad (1.9)$$

where $|k|$ denotes the l^2 -norm of the so-called wave-vector k .

The variational approximations of (1.1) then consists in solving

$$I_\delta = \inf \left\{ \mathcal{E}(v_\delta), v_\delta \in X_\delta, \int_\Omega v_\delta^2 = 1 \right\}. \quad (1.10)$$

Problem (1.10) has at least one minimizer u_δ such that $(u, u_\delta)_{L^2} \geq 0$, which satisfies

$$\forall v_\delta \in X_\delta, \quad \langle \mathcal{A}_{u_\delta} u_\delta, v_\delta \rangle_{X', X} = \lambda_\delta (u_\delta, v_\delta)_{L^2}, \quad (1.11)$$

for some $\lambda_\delta \in \mathbb{R}$. It is easily seen that (see e.g., [7, 63])

$$\lim_{\delta \rightarrow 0} \|u - u_\delta\|_X = 0, \quad (1.12)$$

or, in words, that the approximate ground state eigenfunction converges to the exact ground state eigenfunction in $H_{\#}^1$ -norm, from which we deduce that I_δ and λ_δ converge to I and λ , respectively, when δ goes to 0. Optimal convergence rates have been obtained in [7] (under stronger assumptions on the nonlinearity G) for spectral Fourier discretizations and also for finite element discretizations. First, under suitable and realistic hypotheses on \mathcal{A} , there holds

$$\|u - u_\delta\|_X \lesssim \min_{v_\delta \in X_\delta} \|u - v_\delta\|_X, \quad (1.13)$$

where we denote by $a \lesssim b$ the inequality $a \leq Cb$, with $C > 0$ a constant that is independent of the discretization parameter δ . Also, the eigenvalues converge much faster, similarly as in the linear case where the eigenvalues converge quadratically compared to the eigenvectors, i.e.

$$|\lambda_\delta - \lambda| \lesssim \|u - u_\delta\|_X^2, \quad \text{for } \mu = 0. \quad (1.14)$$

In the nonlinear case, there is an additional term involving an $L^{6/(5-2q)}$ -norm which can be absorbed in the $\|u - u_\delta\|_X^2$ -term under additional regularity assumption that we do not detail here (see e.g. [7, Remark 3] for a precise comment about this). Moreover, the L^2 -norm of the error $u - u_\delta$ as well as its negative Sobolev norms converge faster than the $H_{\#}^1$ -norm of the error. Finally, the dual norm of the residual behaves like the $H_{\#}^1$ -norm of the error, i.e.

$$\|u - u_\delta\|_X \lesssim \|\mathcal{A}_{u_\delta} u_\delta - \lambda_\delta u_\delta\|_{X'} \lesssim \|u - u_\delta\|_X.$$

The first inequality is not trivial as the problem is an eigenvalue problem, but it was shown for the Gross–Pitaevskii equation in [25], and for the Laplace eigenvalue problem in [13].

The outline of this article is the following. In Section 2, we present two equivalent ways of approximating the error between the exact and approximate solutions, the first one relying on the minimization problem (1.1), the second on the eigenvalue problem (1.6). We then present different post-processing methods relying on the presented reconstruction of the error, and approximations thereof. In Section 3, we show that this reconstructed error also appears in a *posteriori* error estimations proposed for this problem, and compare different contributions, namely exposing the main features and range of applicabilities, such as considered models and discretization methods.

2 Two approaches for a derivation of the reconstructed error

In this section, we present two different ways of estimating the error in this context. One is based on the energy minimization problem, and consists in looking at the second-order Taylor expansion of the energy functional. The second one relies on a first-order Taylor expansion of the nonlinear eigenvalue problem. Since the nonlinear eigenvalue problem corresponds to the first-order Euler–Lagrange equations of the minimization problem, it is natural that the two approaches are ultimately equivalent. However, due to the nonlinear structure of the equations and the norm constraint on the solution vector, the derivation is not straightforward, and we present the two derivations to highlight their similarities and differences, as well as pointing out the different references following these schemes.

Note that the equations obtained below for the reconstructed error are similar to what is presented in [54] for a generic equation (without constraints). Also, in [10], such Taylor expansion and first-order error reconstruction is also proposed for the computation of ground state energies in planewave electronic structure calculations for materials systems, involving several eigenvalues.

2.1 Approach based on the energy minimisation problem

In this approach, presented and analyzed in [46] on the Hartree–Fock problem, the initial idea is to provide a lower and an upper bound to the ground state energy $\mathcal{E}(u)$ from the knowledge of $\mathcal{E}(u_\delta)$. Of course, due to the variational statement of the conforming discretization, the following upper bound is classical

$$\mathcal{E}(u_\delta) \geq \mathcal{E}(u).$$

To get a lower bound, the idea is to consider the second-order Taylor expansion of the energy. Before doing so, let us introduce a notation: for any $v \in X$,

$$\Lambda_v = \langle \mathcal{A}_v(v), v \rangle_{X', X},$$

so that the smallest eigenvalue in (1.7) satisfies $\lambda = \Lambda_u$. We then introduce the Lagrangian of the problem defined for $v \in X$, $\nu \in \mathbb{R}$, by

$$\mathcal{L}(v, \nu) = \mathcal{E}(v) - \nu \left(\int_{\Omega} v^2 - 1 \right),$$

and define $\mathcal{E}^w(v)$ as being the value of the Lagrangian at some Λ_w , i.e.

$$\mathcal{E}^w(v) = \mathcal{L}(v, \Lambda_w) = \mathcal{E}(v) - \Lambda_w \left(\int_{\Omega} v^2 - 1 \right).$$

For $w \in X$, denoting by $D_v \mathcal{E}^w$ the differential of $v \in X \mapsto \mathcal{E}^w(v)$, there holds

$$\forall w \in X, \quad \langle [D_v \mathcal{E}^w](u), w \rangle_{X', X} = 0, \quad (2.1)$$

and if u_δ is a solution to the discrete problem (1.11), then

$$\forall w_\delta \in X_\delta, \quad \langle [D_v \mathcal{E}^{u_\delta}](u_\delta), w_\delta \rangle_{X', X} = 0. \quad (2.2)$$

Writing the second-order Taylor expansion of the difference in energies between the two minima over X and over X_δ gives

$$\begin{aligned} \mathcal{E}(u_\delta) - \mathcal{E}(u) &= \mathcal{E}^{u_\delta}(u_\delta) - \mathcal{E}^{u_\delta}(u) \\ &= \langle [D_v \mathcal{E}^{u_\delta}](u_\delta), u_\delta - u \rangle_{X', X} - \frac{1}{2} \langle [D_v^2 \mathcal{E}^{u_\delta}](u_\delta)(u_\delta - u), u_\delta - u \rangle_{X', X} + o(\|u - u_\delta\|_X^2). \end{aligned}$$

Noting that $(u - u_\delta, u_\delta)_{L^2} = -\frac{1}{2}\|u_\delta - u\|_{L^2}^2$, and defining $\varepsilon_0 = \|u_\delta - u\|_{L^2}$ we can state

$$u - u_\delta = -\frac{1}{2}\varepsilon_0^2 u_\delta + w, \quad \text{with } w \perp u_\delta,$$

where $\|w\|_X \simeq \varepsilon = \|u_\delta - u\|_X$, and of course $\varepsilon_0 \leq \varepsilon$. Using the above decomposition and (2.2), the energy difference can be written as

$$\begin{aligned} \mathcal{E}(u_\delta) - \mathcal{E}(u) &= \langle [D_v \mathcal{E}^{u_\delta}](u_\delta), \frac{1}{2}\varepsilon_0^2 u_\delta - w \rangle_{X', X} \\ &\quad - \frac{1}{2} \langle [D_v^2 \mathcal{E}^{u_\delta}](u_\delta) \left(\frac{1}{2}\varepsilon_0^2 u_\delta - w \right), \frac{1}{2}\varepsilon_0^2 u_\delta - w \rangle_{X', X} + o(\varepsilon^2) \\ &= -\langle [D_v \mathcal{E}^{u_\delta}](u_\delta), w \rangle_{X', X} - \frac{1}{2} \langle [D_v^2 \mathcal{E}^{u_\delta}](u_\delta)(w), w \rangle_{X', X} + o(\varepsilon^2). \end{aligned}$$

Introducing the bilinear form $a_v \equiv [D_v^2 \mathcal{E}^v](v)(\cdot, \cdot) = \langle [D_v^2 \mathcal{E}](v)(\cdot, \cdot) \rangle_{X', X} - \Lambda_v \langle \cdot, \cdot \rangle_{X', X}$, the previous equation reads

$$\mathcal{E}(u_\delta) - \mathcal{E}(u) = -\langle [D_v \mathcal{E}^{u_\delta}](u_\delta), w \rangle_{X', X} - \frac{1}{2} a_{u_\delta}(w, w) + o(\varepsilon^2). \quad (2.3)$$

Using the fact that the lower eigenvalue is simple, that is there is a gap $\lambda_2 - \lambda > 0$ between the two first eigenvalues of \mathcal{A}_u , see e.g. Lemma 1 in [7], there holds:

Proposition 2.1. *There exists a constant $c_u > 0$ such that, for any $v \in X$, $v \perp u$,*

$$a_u(v, v) \geq c_u \|v\|_X^2.$$

From this result, we obtain the following proposition.

Proposition 2.2. *Assume that $\varepsilon := \|u - u_\delta\|_X$ is small enough, then, there exist a constant $c_{u_\delta} > 0$ such that, for any $v \in X$, $v \perp u_\delta$*

$$a_{u_\delta}(v, v) \geq c_{u_\delta} \|v\|_X^2.$$

Using this last proposition, we derive that there exists a unique solution $\hat{w} \in X^\perp$ (where X^\perp is the orthogonal to u_δ in X) called the *reconstructed error* such that

$$a_{u_\delta}(\hat{w}, \psi) = -\langle [D_v \mathcal{E}^{u_\delta}](u_\delta), \psi \rangle_{X', X} \quad \forall \psi \in X \quad (2.4)$$

which allows to rewrite (2.3) as

$$\mathcal{E}(u) = \mathcal{E}(u_\delta) - \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w}) + \frac{1}{2}a_{u_\delta}(w - \hat{w}, w - \hat{w}) + o(\varepsilon^2) \quad (2.5)$$

and yields the authors in [46] to the inequality

$$\mathcal{E}(u) \geq \mathcal{E}(u_\delta) - \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w}) + o(\varepsilon^2), \quad (2.6)$$

so that $\mathcal{E}(u_\delta) - \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w})$ is an asymptotic lower bound to the exact energy $\mathcal{E}(u)$.

To compute this reconstructed error and the associated lower bound we have to solve problem (2.4) which, of course, cannot be done exactly and thus needs to be discretized in a larger space X_{δ_f} than X_δ . This problem reads: Find $\hat{w}_{\delta_f} \in X_{\delta_f}^\perp$ (where $X_{\delta_f}^\perp$ is the orthogonal to u_δ in X_{δ_f}) such that

$$\langle (\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)\hat{w}_{\delta_f}, \psi_{\delta_f} \rangle_{X', X} = -\langle (\mathcal{A} + g(u_\delta^2) - \lambda_\delta)u_\delta, \psi_{\delta_f} \rangle_{X', X} \quad \forall \psi_{\delta_f} \in X_{\delta_f}^\perp. \quad (2.7)$$

As we shall see later this can be further refined by noticing that w and \hat{w} – or rather its discrete representation \hat{w}_{δ_f} – are ε^2 close. Indeed, we first note that

$$[D_v \mathcal{E}^{u_\delta}](u) = [D_v \mathcal{E}^{u_\delta}](u_\delta) + [D_v^2 \mathcal{E}^{u_\delta}](u_\delta)(u - u_\delta) + o(\varepsilon).$$

Since

$$[D_v \mathcal{E}^{u_\delta}](u) = [D_v \mathcal{E}^u](u) + (\lambda_\delta - \lambda)u,$$

we show using (2.1) and (1.14) that

$$[D_v \mathcal{E}^{u_\delta}](u_\delta) + [D_v^2 \mathcal{E}^{u_\delta}](u_\delta)(u - u_\delta) = o(\varepsilon). \quad (2.8)$$

Combining (2.4) and (2.8), we deduce that for any $\psi \in X$

$$a_{u_\delta}(u - u_\delta - \hat{w}, \psi) = o(\varepsilon)\|\psi\|_X. \quad (2.9)$$

Since $\|u - u_\delta - w\|_X = O(\varepsilon^2)$, there holds

$$a_{u_\delta}(w - \hat{w}, \psi) = o(\varepsilon)\|\psi\|_X,$$

and Proposition 2.1 together with (2.1) show that

$$\|w - \hat{w}\|_X = o(\varepsilon). \quad (2.10)$$

Using (2.5), we first deduce that (2.6) can be improved as an equality

$$\mathcal{E}(u) = \mathcal{E}(u_\delta) - \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w}) + o(\varepsilon^2),$$

and second that $u_\delta + \hat{w}$ is a better approximation to u than u_δ . However $u_\delta + \hat{w}$ is yet not of norm 1, in order to cure this it remains to tune α^* such that $\hat{u}_\delta = (1 - \alpha^*)u_\delta + \hat{w}$ is of norm 1 in $L^2(\Omega)$ (which is always possible since $\|\hat{w}\|_X = O(\varepsilon^2)$)². Note that \hat{u}_δ is computable only from the knowledge of u_δ by inverting the linear problem (2.7) on a finer grid.

With this we obtain a quadratic approximation both in X -norm and in energy. Following [46], we write

Theorem 2.1. *Let us assume that $\|u - u_\delta\|_X$ is small enough, then $\hat{u}_\delta = (1 - \alpha^*)u_\delta + \hat{w}$ verifies*

$$\|u - \hat{u}_\delta\|_X = o(\|u - u_\delta\|_X).$$

In addition, if $G \in C^3((0, \infty), \mathbb{R})$ then

$$\|u - \hat{u}_\delta\|_X \lesssim \|u - u_\delta\|_X^2,$$

and

$$|\mathcal{E}(u) - \mathcal{E}(\hat{u}_\delta)| \lesssim \|u - u_\delta\|_X^4.$$

The two last improved estimates follow directly from the previous analysis by changing the $o(\varepsilon^q)$ by $\mathcal{O}(\varepsilon^{q+1})$ with $q = 0, 1$, or 2 .

Remark 2.1. *The quantities \hat{u}_δ and \hat{w} , which can be computed with the knowledge of u_δ can be used in practice for two different (complementary) goals. First, they can be used for a refined approximation of the solution, saying that \hat{u}_δ is a better approximation to the exact solution u than u_δ , and similarly for the energy using \hat{w} , that is $\mathcal{E}(u) \simeq \mathcal{E}(u_\delta) - \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w})$. Second, they can be used to obtain a refined error bound, as $|\mathcal{E}(u) - \mathcal{E}(u_\delta)| \simeq \frac{1}{2}a_{u_\delta}(\hat{w}, \hat{w})$. These two possible approaches will be detailed separately in Sections 2.3 and 3.*

Remark 2.2. *Let us mention a few works which consider the nonlinear eigenvalue problem of this type with the angle of the energy minimization, namely [39, 26, 34, 32, 62]. In [32], the goal is to provide an adaptive procedure for the computation of the solutions of this problem, by minimizing the energy directly, by combining gradient flow iterations and adaptive finite element mesh refinements. In these other works, proofs for the exponential convergence of the continuous Sobolev gradient flow are provided [39, 34], respectively for the discrete gradient flow in [26, 62]. Note however that these works do not provide specific error bounds on the computed solutions.*

2.2 Approach based on the nonlinear eigenvalue problem

In the previous section, we derived the reconstructed error equation (2.7) starting from the energy minimization problem (1.1). In this section, we aim at arriving at the same equation starting from the eigenvalue problem (1.6), which reads: find (u, λ) such that $\|u\|_{L^2} = 1$ and

$$\mathcal{F}(u, \lambda) := \mathcal{A}_u(u) - \lambda u = 0. \tag{2.11}$$

In order to improve $(u_\delta, \lambda_\delta)$, one can think of using a single step of a Newton method in a finer discrete space X_{δ_f} as introduced in the previous section that reads: find $\tau_{\delta_f} \in X_{\delta_f}$, $(\tau_{\delta_f}, u_\delta)_{L^2} = 0$

²note that another normalisation can be obtained by setting $\hat{u}_\delta = \beta^*(u_\delta + \hat{w})$

and $\gamma_{\delta_f} \in \mathbb{R}$ such that

$$\langle [D_v \mathcal{F}](u_\delta, \lambda_\delta)(\tau_{\delta_f}), \psi_{\delta_f} \rangle_{X', X} + \langle [D_\mu \mathcal{F}](u_\delta, \lambda_\delta)(\gamma_{\delta_f}), \psi_{\delta_f} \rangle_{X', X} = \langle \mathcal{F}(u_\delta, \lambda_\delta), \psi_{\delta_f} \rangle_{X', X} \quad \forall \psi_{\delta_f} \in X_{\delta_f}. \quad (2.12)$$

Note that the solution τ_{δ_f} is searched in the $L^2(\Omega)$ -orthogonal complement of u_δ , instead of X , since the problem is not well-posed on X_δ , hence not well-conditioned on X . This can also be linked to the normalization constraint $\|u_\delta\|_{L^2}^2 = 1$. Indeed, this condition guarantees that the first-order equation relative to the constraint is satisfied. Since we have

$$\forall v \in X, \quad \forall \mu \in \mathbb{R}, \quad \mathcal{F}(v, \mu) = (\mathcal{A} + g(v^2) - \mu)v,$$

and the differential $D_v \mathcal{F}(u_\delta, \lambda_\delta)$ of \mathcal{F} at $(u_\delta, \lambda_\delta)$ writes

$$\forall v \in X, \quad D_v \mathcal{F}(u_\delta, \lambda_\delta)(v) = (\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)v,$$

the single step of the Newton method in the fine grid reads in a strong form as: Find $\tau_{\delta_f} \in X_{\delta_f}$, $(\tau_{\delta_f}, u_\delta)_{L^2} = 0$ such that

$$(\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)\tau_{\delta_f} = -(\mathcal{A} + g(u_\delta^2) - \lambda_\delta)u_\delta \quad \text{in } X_{\delta_f}^\perp, \quad (2.13)$$

which is similar to (2.7). We thus propose a norm 1 improved approximation of u_δ as

$$\tilde{u}_\delta = (1 - \alpha')u_\delta + \tau_{\delta_f},$$

for some $\alpha' \in \mathbb{R}$. Then, the difference between the post-processed energy and the approximate energy allows to estimate the error between the exact energy and the approximate one, that is

$$\mathcal{E}(u) - \mathcal{E}(u_\delta) \simeq \mathcal{E}(\tilde{u}_\delta) - \mathcal{E}(u_\delta).$$

Compared to the previous *a posteriori* estimation (2.6), the bounds here are not guaranteed. Nevertheless, they asymptotically match the true error, provided that the space X_{δ_f} is large enough.

Remark 2.3. *The proposed procedure in fact corresponds to a standard Newton step on the modified \mathcal{F} functional including the normalization condition $\|u_\delta\|_{L^2}^2 = 1$ as defined below in (3.1). Hence, the quadratic convergence results of the Newton method apply in this context, so that this proposed post-processing doubles the convergence rate of the X -norm of the eigenfunctions and also of the energy.*

2.3 Practical approximations of the reconstructed error

The reconstructed error presented in (2.7) gives a first-order approximation of the discretization error for problem (1.7). It appears that several contributions on post-processing and error estimation for nonlinear eigenvalue problems [46, 16, 25, 11, 6, 12] are based on this reconstructed error, and approximations thereof. We try to give an overview of these methods and compare them in this section.

The main focus of this article is on nonlinear eigenvalue problems. But naturally, post-processing methods were first developed for linear eigenvalue problems – that are themselves actually nonlinear problems – before nonlinear eigenvalue problems. Therefore we first describe a few methods that

were developed for linear eigenvalue problems. Already in 1999 by Xu and Zhou in [59], a two-grid method has been proposed to efficiently solve eigenvalue problems. It consists in first solving an eigenvalue problem on a coarse finite element mesh, and then solving a boundary value problem on a fine mesh, in order to improve the eigenvector, then an improved eigenvalue is obtained through a Rayleigh quotient. This avoids paying the full price of solving the eigenvalue problem on the fine mesh. Later on, in [51], another post-processing method was presented, this time proposing only an improved eigenvalue, similarly requiring the resolution of a linear boundary value problem.

Other works include multigrid methods, such as [49, 44, 31], where the idea is to first solve an eigenvalue problem on a very coarse mesh, and then to have a family of meshes, and improve the initial solution

- either by solving several linear problems on finer and finer meshes [44],
- or applying some gradient recovery operator [31],

correcting this way the initially found eigenvector and eigenvalue.

In the same spirit as the two-grid and multi-grid post-processing methods for linear eigenvalue problems, several methods have been proposed for nonlinear eigenvalue problems. To show how these methods relate to the reconstructed error equation (2.13), let us first express this equation in terms of the post-processed solution \tilde{u}_δ and not only the correction τ_{δ_f} . From (2.13) there holds

$$(\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)(u_\delta + \tau_{\delta_f}) = (\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)u_\delta - (\mathcal{A} + g(u_\delta^2) - \lambda_\delta)u_\delta,$$

hence

$$(\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)(u_\delta + \tau_{\delta_f}) = 2[g'(u_\delta^2)u_\delta^2]u_\delta.$$

The problem posed on the fine grid corresponding to the reconstructed error is therefore: Find $u_{\delta_f} \in X_{\delta_f}$ such that

$$(\mathcal{A} + 2g'(u_\delta^2)u_\delta^2 + g(u_\delta^2) - \lambda_\delta)u_{\delta_f} = 2[g'(u_\delta^2)u_\delta^2]u_\delta. \quad (2.14)$$

Related to this post-processing, a two-level discretization technique has been presented in [6] where the authors propose three different schemes for the Gross–Pitaevskii equation. All start by solving the nonlinear eigenvalue problem in a small basis. Then three alternatives are proposed. The first one (scheme 1) is to solve a linear eigenvalue problem on the large basis set, fixing the nonlinearity with the coarse solution. The second one (scheme 2a) consists in solving the following boundary value problem

$$(\mathcal{A} + g(u_\delta^2))u_{\delta_f} = \lambda_\delta u_\delta \quad \text{in } X_{\delta_f},$$

where the term $g(u_\delta^2)$ remains on the left hand side. The third scheme (called 2b) amounts to solving a linear boundary value problem on the fine space, putting the nonlinear term on the right hand side, namely solving

$$\mathcal{A}u_{\delta_f} = \lambda_\delta u_\delta - g(u_\delta^2)u_\delta \quad \text{in } X_{\delta_f}. \quad (2.15)$$

The numerical analysis of the first scheme gives the following estimates for a plane wave approximation :

$$\begin{aligned} \|u - u_{\delta_f}\|_X &\lesssim M^{-2} \|u - u_\delta\|_X, \\ |\lambda - \lambda_{\delta_f}| + \|u - u_{\delta_f}\|_{L^2} &\lesssim \|u - u_\delta\|_X^2, \\ |E(u) - E(u_{\delta_f})| &\lesssim M^{-4} |E(u) - E(u_\delta)|. \end{aligned}$$

Related to this post-processing, a two-level discretization technique was also proposed earlier in [33] in the finite element context in the form of (2.15) based on the use of quasi-orthogonality properties of a Clement type operator, hence difficult to transpose to the planewave method. They were able to improve not only the approximation of the eigenfunctions but also the eigenvalues. The same approach based on the Localized Orthogonal Decomposition (LOD) has been recently extended to treat both the time-dependent Gross–Pitaevskii equation and the nonlinear Gross–Pitaevskii eigenvalue problem [23].

Another work proposing a post-processing technique for nonlinear eigenvalue problems is [12], see [13, 24] for the proofs of the estimates. This method is based on a perturbative expansion of the eigenvalues and eigenvectors in order to post-process the eigenfunctions and the energy at a very low computational cost. This method was presented in the case of the Kohn–Sham equations, which is a nonlinear eigenvalue problem, where one needs to compute a few low-lying eigenstates of the considered nonlinear operator. The method consists in solving first the full eigenvalue problem on a small planewave space, and then to post-process the eigenvectors and eigenvalues on a larger planewave using the derived perturbation expansion. This method particularly exploits the diagonal structure of the Laplace operator when expressed in planewaves, which makes the post-processing particularly cheap to perform, but makes the method difficult to generalize to different types of discretizations. Translated on our one-eigenpair nonlinear problem, the linear boundary value problem solved on the large discretization space aims at computing $\tau_{\delta_f} \in X_{\delta_f}$ solution to

$$(-\Delta - \lambda_\delta)\tau_{\delta_f} = -(\mathcal{A} + g(u_\delta^2) - \lambda_\delta)u_\delta, \quad \text{in } X_{\delta_f}, \quad (2.16)$$

in a case where the operator $(-\Delta - \lambda_\delta)$ is diagonal, thus only two FFTs per eigenvalue are required to compute the residual in a fine grid. Compared to (2.13), some terms in the operator on the left hand side are removed. They are actually shown to be asymptotically small compared to the Laplace operator. In terms of errors, one obtains that the eigenvectors and energy is improved by a factor M^{-2} . So the improvement of this perturbation method is limited to M^{-2} for the eigenvectors or M^{-4} for the energy, whereas, in the two-grid case, the convergence rate of the eigenfunctions can be doubled. Nevertheless it is much less expensive.

Beyond two-grid methods, there also exist multi-grid methods which use more than two grids for computing an approximation of the solution on a fine basis. For nonlinear eigenvalue problems, several of them have been proposed, such as [38, 37, 58]. In these three contributions, the idea, similarly as in the two-grid case, is to first solve a nonlinear eigenvalue problem on a small discretization space, and then to post-process the solution. In the multi-grid case, this post-processing consists of several steps, often including the resolution of boundary value problems on spaces of larger and larger size, or amounting to use a multigrid technique on the larger grids, in order to even avoid solving boundary value problems on those grids. In that case, the resolution of the problems cannot be directly linked to equation (2.14), except at the first level.

3 Using the reconstructed error for *a posteriori* error estimation

The previous section, by proposing, as in [46], a better approximation \tilde{u}_δ of the exact solution u using the reconstructed error equation, allows to estimate the error $u - u_\delta$ by the difference between the approximate solution u_δ and the post-processed solution, i.e.

$$u - u_\delta \simeq \tilde{u}_\delta - u_\delta.$$

This approach is in principle valid for any type of conforming discretization.

We will see below that several works use such a post-processing step to estimate the error in the context of eigenvalue problems, sometimes in a modified way in order to obtain guaranteed error bounds. In the generic context of nonlinear problems, the contribution [54] presents an elegant way to combine a post-processing step with guaranteed error bounds. Namely, they consider a so-called “nonsplit residual” which corresponds to the reconstructed error $[D\mathcal{F}(u, \lambda)]^{-1}(\mathcal{F}(u, \lambda))$ introduced in (2.12) where $\mathcal{F} : X \times \mathbb{R} \rightarrow X \times \mathbb{R}$ is defined by

$$\mathcal{F}(u, \lambda) := \begin{cases} \mathcal{A}_u(u) - \lambda u \\ \int_{\Omega} u^2 - 1, \end{cases} \quad (3.1)$$

(with a slight change in notation for \mathcal{F} with respect to (2.11)) and $[D\mathcal{F}(u, \lambda)] : X \times \mathbb{R} \rightarrow X' \times \mathbb{R}$ denotes the differential of \mathcal{F} at point (u, λ) . Then (see [54, Theorem 1]) considering the stability constant

$$\gamma(u, \lambda) = \|[D\mathcal{F}(u, \lambda)]^{-1}\|_{\mathcal{L}(X' \times \mathbb{R}, X \times \mathbb{R})},$$

and a local nonlinearity indicator

$$L(\alpha) = \sup_{(v, \mu) \in \bar{B}_{\alpha}(u, \lambda)} \|D\mathcal{F}(v, \mu) - D\mathcal{F}(u, \lambda)\|_{\mathcal{L}(X' \times \mathbb{R}, X \times \mathbb{R})},$$

if the validity criterion

$$\alpha(u, \lambda) := 2\gamma(u, \lambda)L(2\varepsilon(u, \lambda)) \leq 1$$

is satisfied, with $\varepsilon(u, \lambda) = \|[D\mathcal{F}(u, \lambda)]^{-1}(\mathcal{F}(u, \lambda))\|$ then problem $\mathcal{F}(v, \mu) = 0$ has a unique solution $(u^*, \lambda^*) \in \bar{B}_{2\varepsilon(u, \lambda)}(u, \lambda)$ and the error is bounded by

$$\|(u^*, \lambda^*) - (u, \lambda)\|_Y \leq 2\varepsilon(u, \lambda).$$

Therefore, this gives a generic way to estimate the error, at the price of estimating the size of $\varepsilon(u, \lambda)$ and not exactly solving the reconstructed error equation.

Regarding linear eigenvalue problems, several *a posteriori* error estimations have been proposed, including [41, 28, 42, 36, 48, 35, 29, 10, 45] and references therein. Often, the estimations are based on a post-processing step, which allows to obtain a more accurate solution to the problem, more or less directly related to the error bound.

Concerning nonlinear eigenvalue problems, there exist several works proposing error estimates for the problem of interest in this article. Error bounds were proposed in [19], an article presenting adaptive refinement techniques for finite element simulations of Gross–Pitaevskii type equations. This work was later extended to the finite element simulations of Kohn–Sham equations in [16]. In these two works, the error is proved to be asymptotically bounded by an error indicator (up to a constant) involving the computation L^2 -norm of the residual. However, the error estimates are not fully guaranteed.

In [25] we provided an *a posteriori* error estimation for the Gross–Pitaevskii equation, based on a careful two-steps approximation. A first coarse bound based on the analysis of the first-order Taylor expansion of the error (see Section 2.2) allows to characterize the asymptotic regime and to validate when the second proposed error bound, which is close to the real error, is guaranteed.

Therefore conditions guaranteeing that the second bound is valid can be checked in practice. More precisely, the first bound is based on Newton–Kantorovich theorem (see e.g. [4]) which similarly as in [54] provides conditions ensuring that there exists an exact solution in the vicinity of the approximate solution, and that the error between the exact and approximate solution is bounded by

$$\|u - u_\delta\|_X + |\lambda - \lambda_\delta| \leq 2\|[D\mathcal{F}(u_\delta, \lambda_\delta)]^{-1}\|_{(X',\mathbb{R}), (X,\mathbb{R})}\|\mathcal{F}(u_\delta, \lambda_\delta)\|_{(X',\mathbb{R})}.$$

The factor 2 in this estimation allows to absorb, in a guaranteed way, the higher-order terms of the Taylor development. An important part of the contribution consists of showing that the differential $D\mathcal{F}$ is invertible at $(u_\delta, \lambda_\delta)$ and to bound the norm of its inverse. Indeed, the main part of $[D\mathcal{F}(u_\delta, \lambda_\delta)]^{-1}$ is $(\Delta)^{-1}$, that is an isometry between X' and X , and the remaining part in $[D\mathcal{F}(u_\delta, \lambda_\delta)]^{-1}$ is of lower order in terms of differential operator. In the second refined bound presented in this article, the term $[D\mathcal{F}(u_\delta, \lambda_\delta)]^{-1}\mathcal{F}(u_\delta, \lambda_\delta)$ is estimated from $\Delta^{-1}\mathcal{F}(u_\delta, \lambda_\delta)$ plus complementary terms that are shown to be negligible thanks to the first bound. Asymptotically, the following bound is obtained

$$\|u - u_\delta\|_X \leq \alpha\|\mathcal{F}(u_\delta, \lambda_\delta)\|_{(X',\mathbb{R})},$$

where α can be taken as close to 1 as we wish when the discretization parameter δ is refined, and the second component of the residual $\mathcal{F}(u_\delta, \lambda_\delta)$ is zero since the norm constraint is exactly satisfied. The main drawback of this method is the high computational cost to obtain these bounds. Indeed, to check that the necessary assumptions are satisfied, a linear eigenvalue problem in the discrete space has to be solved, for which the lowest two eigenvalues have to be computed. Note that [25] considers planewave discretization, but, in opposition to the perturbation approach, it can be generalized to other discretization methods such as finite elements. Note finally that the above estimate is further pushed to separate the two sources of error when the method is actually implemented and an iterative algorithm is resorted to solving the resulting nonlinear discrete problem : the discretization error and the iteration error.

Another method has been proposed more recently for the Kohn–Sham problem [10] for accurately estimating the error for quantities of interest. This work also makes use of the reconstructed error equation. Moreover the proposed bounds are computable and accurate, in the sense that they are close to the real error. They involve a post-processing step that is quite cheap, and similar to the one of [12], i.e. they only need a few FFTs on the fine space.

In the paper [57], an *a posteriori* estimation for a finite element discretization was provided. The resolution of an auxiliary boundary value problem is also necessary to obtain computable upper bound of the error. The results are asymptotic in the sense that they are valid for sufficiently small meshes, but without an *a posteriori* guarantee that the mesh is fine enough so that the bounds are valid. Note that, in the frame of finite element methods, local estimators based on the use of Prager–Synge techniques allow to propose alternative global bounds that can be used to improve the precision by locally refining the finite element mesh [17, 16, 13].

4 Conclusion

In this article, we showed how the first-order Taylor expansion of the solved equation is related to post-processing methods as well as error estimation techniques, and summarized several works based

on this especially for nonlinear eigenvalue problems. To complete the set of examples presented above, we selected a few and compare them in a unified way in Table 1. Namely, we compare the problems that are originally considered, the goals of the different contributions, as well as the computational cost of the procedure.

Method	Equation	Number of eigenvalues	Discretization method	Generic type of estimates	Cost
Nonlinear two-grid method [33]	$(-\Delta + V + g(u^2))u = \lambda u$	1 lowest	Finite Elements (FE)	$\ u - u_{\delta_f}\ _{H^1} \lesssim \delta^3 + \ u - u_{\delta}\ _{H^1}$ $ \lambda - \lambda_{\delta_f} + \ u - u_{\delta_f}\ _{L^2} \lesssim \delta^4 + \delta^2 \ u - u_{\delta_f}\ _{H^1}$	Boundary Value Problem (BVP)
Nonlinear two-grid method [6]	$(-\Delta + V + g(u^2))u = \lambda u$	1 lowest	FE or planewaves (PW)	(in FE) if $\ u - u_{\delta}\ _{H^1} \lesssim \delta^{-\sigma}$, then $\ u - u_{\delta_f}\ _{H^1} \lesssim \delta^{-\sigma-2} + \delta_f^{-\sigma}$ $ E(u) - E(u_{\delta_f}) \lesssim \delta^{-2\sigma-4} + \delta_f^{-2\sigma}$	BVP or linear eigenvalue problem
Nonlinear perturbation method [12]	$(-\Delta + V + V_{\rho_{\Phi}})\phi_i = \lambda_i \phi_i$ with $\Phi = (\phi_1, \dots, \phi_K)$.	K lowest	PW	if N_c cutoff in momentum space, $\ u - u_{\delta_f}\ _{H^1} \lesssim M^{-2} \ u - u_{\delta}\ _{H^1}$ Energy: $ E(u) - E(u_{\delta_f}) \lesssim M^{-2} E(u) - E(u_{\delta}) $	Residual computation
<i>A posteriori</i> estimation for Hartree–Fock [46]	$(-\Delta + V + (\rho_{\Phi} \star \frac{1}{ x }))\phi_i = \lambda_i \phi_i$, with $\Phi = (\phi_1, \dots, \phi_K)$.	K lowest	Any	$\ u - u_{\delta_f}\ _{H^1} \lesssim \ u - u_{\delta}\ _{H^1}^2$ $ E(u) - E(u_{\delta_f}) \lesssim E(u) - E(u_{\delta}) ^2$	BVP
<i>A posteriori</i> estimation for Gross–Pitaevskii [25]	$(-\Delta + V + g(u^2))u = \lambda u$,	1 lowest	PW	$\ u - u_{\delta}\ _X \lesssim \ \mathcal{F}(u_{\delta}, \lambda_{\delta})\ _{X'}$	BVP
<i>A posteriori</i> estimation for nonlinear problem [54]	$F(u) = 0$	no eigenvalue	Any	$\ u - u_{\delta}\ _X \lesssim \ \mathcal{F}(u_{\delta}, \lambda_{\delta})\ _{X'}$	BVP
<i>A posteriori</i> estimation for Kohn–Sham [10]	$(-\Delta + V_{\rho_{\Phi}})\phi_i = \lambda_i \phi_i$, with $\Phi = (\phi_1, \dots, \phi_K)$.	K lowest	PW	$\ \gamma_{\Phi} - \gamma_{\Phi_{\delta_f}}\ _X \lesssim \ \mathcal{F}(\Phi)\ _{X'}$ (density matrix)	BVP

Table 1: Comparative table of different post-processing and error estimation methods.

Dedication

We dedicate this article to the late Professor Roland Glowinski who was an inspiration to so many generations of applied and interdisciplinary mathematicians. The second author (YM) had the chance to follow his DEA (then master) courses at the Université Pierre et Marie Curie (then Sorbonne Université) and to benefit from his inspiring teachings both in these courses and in the lectures he was given. Even if the nonlinear eigenvalue problems he was interested in [30] were of a different nature, applications to molecular dynamics and electronic structure calculation [60, 61] has been of interest for him for many years and will undoubtedly benefit from his ideas for a long time to come.

Funding

For this project GD is supported by the French “Investissements d’Avenir” program, project ISITE-BFC (contract ANR-15-IDEX-0003) and YM is supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 810367), project EMC2.

References

- [1] R. ALTMANN, P. HENNING, AND D. PETERSEIM, *The j -method for the Gross–Pitaevskii eigenvalue problem*, Numer. Math., 148 (2021), pp. 575–610.
- [2] W. BAO AND W. TANG, *Ground-state solution of Bose–Einstein condensate by directly minimizing the energy functional*, J. Comput. Phys., 187 (2003), pp. 230–254.
- [3] M. CALIARI, A. OSTERMANN, S. RAINER, AND M. THALHAMMER, *A minimisation approach for computing the ground state of Gross–Pitaevskii systems*, J. Comput. Phys., 228 (2009), pp. 349–360.
- [4] G. CALOZ AND J. RAPPAZ, *Numerical analysis for nonlinear and bifurcation problems*, Handb. Numer. Anal., 5 (1997), pp. 487–637.
- [5] E. CANCÈS, *SCF algorithms for HF electronic calculations*, in Mathematical Models and Methods for Ab Initio Quantum Chemistry, M. Defranceschi and C. Le Bris, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2000, pp. 17–43.
- [6] E. CANCÈS, R. CHAKIR, L. HE, AND Y. MADAY, *Two-grid methods for a class of nonlinear elliptic eigenvalue problems*, IMA J. Numer. Anal., 38 (2018), pp. 605–645.
- [7] E. CANCÈS, R. CHAKIR, AND Y. MADAY, *Numerical analysis of nonlinear eigenvalue problems*, J. Sci. Comput., 45 (2010), pp. 90–117.
- [8] E. CANCÈS, R. CHAKIR, AND Y. MADAY, *Numerical analysis of the planewave discretization of some orbital-free and Kohn–Sham models*, Esaim Math. Model. Numer. Anal., 46 (2012), pp. 341–388.

- [9] E. CANCÈS, M. DEFRANCESCHI, W. KUTZELNIGG, C. LE BRIS, AND Y. MADAY, *Computational quantum chemistry: A primer*, in Handbook of Numerical Analysis, vol. 10, Elsevier, 2003, pp. 3–270.
- [10] E. CANCÈS, G. DUSSON, G. KEMLIN, AND A. LEVITT, *Practical error bounds for properties in Plane-Wave electronic structure calculations*, SIAM J. Sci. Comput., 44 (2022), pp. B1312–B1340.
- [11] E. CANCÈS, G. DUSSON, Y. MADAY, B. STAMM, AND M. VOHRALÍK, *A perturbation-method-based a posteriori estimator for the planewave discretization of nonlinear schrödinger equations*, C. R. Math., 352 (2014), pp. 941–946.
- [12] ———, *A perturbation-method-based post-processing for the planewave discretization of Kohn–Sham models*, J. Comput. Phys., 307 (2016), pp. 446–459.
- [13] ———, *Guaranteed and robust a posteriori bounds for laplace eigenvalues and eigenvectors: Conforming approximations*, SIAM J. Numer. Anal., 55 (2017), pp. 2228–2254.
- [14] É. CANCÈS, G. KEMLIN, AND A. LEVITT, *Convergence analysis of direct minimization and Self-Consistent iterations*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 243–274.
- [15] E. CANCÈS AND C. LE BRIS, *Can we outperform the DIIS approach for electronic structure calculations?*, Int. J. Quantum Chem., 79 (2000), pp. 82–90.
- [16] H. CHEN, X. DAI, X. GONG, L. HE, AND A. ZHOU, *Adaptive finite element approximations for Kohn–Sham models*, Multiscale Model. Simul., 12 (2014), pp. 1828–1869.
- [17] H. CHEN, X. GONG, L. HE, AND A. ZHOU, *Adaptive finite element approximations for a class of nonlinear eigenvalue problems in quantum physics*, Adv. Appl. Math. Mech., 3 (2011), pp. 493–518.
- [18] H. CHEN, X. GONG, AND A. ZHOU, *Numerical approximations of a nonlinear eigenvalue problem and applications to a density functional model*, Math. Methods Appl. Sci., 33 (2010), pp. 1723–1742.
- [19] H. CHEN, L. HE, AND A. ZHOU, *Finite element approximations of nonlinear eigenvalue problems in quantum physics*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 1846–1865.
- [20] X. DAI, L. HE, AND A. ZHOU, *Convergence and quasi-optimal complexity of adaptive finite element computations for multiple eigenvalues*, IMA J. Numer. Anal., 35 (2015), pp. 1934–1977.
- [21] X. DAI, J. XU, AND A. ZHOU, *Convergence and optimal complexity of adaptive finite element eigenvalue computations*, Numer. Math., 110 (2008), pp. 313–355.
- [22] C. M. DION AND E. CANCÈS, *Ground state of the time-independent Gross–Pitaevskii equation*, Comput. Phys. Commun., 177 (2007), pp. 787–798.

- [23] C. DÖDING, P. HENNING, AND J. WÄRNEGÅRD, *An efficient two level approach for simulating bose-einstein condensates*, arXiv preprint arXiv:2212.07392, (2022).
- [24] G. DUSSON, *Post-processing of the plane-wave approximation of schrödinger equations. part II: Kohn–Sham models*, IMA J. Numer. Anal., 41 (2020), pp. 2456–2487.
- [25] G. DUSSON AND Y. MADAY, *A posteriori analysis of a nonlinear Gross–Pitaevskii-type eigenvalue problem*, IMA J. Numer. Anal., (2016), p. drw001.
- [26] E. FAOU AND T. JÉZÉQUEL, *Convergence of a normalized gradient algorithm for computing ground states*, IMA J. Numer. Anal., 38 (2017), pp. 360–376.
- [27] J. J. GARCÍA-RIPOLL AND V. M. PÉREZ-GARCÍA, *Optimizing schrödinger functionals using sobolev gradients: Applications to quantum mechanics and nonlinear optics*, SIAM J. Sci. Comput., 23 (2001), pp. 1316–1334.
- [28] J. GEDICKE AND C. CARSTENSEN, *A posteriori error estimators for convection–diffusion eigenvalue problems*, Comput. Methods Appl. Mech. Eng., 268 (2014), pp. 160–177.
- [29] S. GIANI, L. GRUBIŠIĆ, H. HAKULA, AND J. S. OVAL, *A posteriori error estimates for elliptic eigenvalue problems using auxiliary subspace techniques*, J. Sci. Comput., 88 (2021).
- [30] R. GLOWINSKI, S. LEUNG, H. LIU, AND J. QIAN, *On the numerical solution of nonlinear eigenvalue problems for the monge-ampère operator*, ESAIM: Control, Optimisation and Calculus of Variations, 26 (2020), p. 118.
- [31] H. GUO, Z. ZHANG, AND R. ZHAO, *Superconvergent two-grid methods for elliptic eigenvalue problems*, J. Sci. Comput., 70 (2017), pp. 125–148.
- [32] P. HEID, B. STAMM, AND T. P. WIHLER, *Gradient flow finite element discretizations with energy-based adaptivity for the Gross-Pitaevskii equation*, J. Comput. Phys., 436 (2021), p. 110165.
- [33] P. HENNING, A. MÅLQVIST, AND D. PETERSEIM, *Two-level discretization techniques for ground state computations of Bose-Einstein condensates*, SIAM J. Numer. Anal., 52 (2014), pp. 1525–1550.
- [34] P. HENNING AND D. PETERSEIM, *Sobolev gradient flow for the Gross–Pitaevskii eigenvalue problem: Global convergence and computational efficiency*, SIAM J. Numer. Anal., 58 (2020), pp. 1744–1772.
- [35] M. F. HERBST, A. LEVITT, AND E. CANCÈS, *A posteriori error estimation for the non-self-consistent Kohn–Sham equations*, Faraday Discuss., 224 (2020), pp. 227–246.
- [36] T. HORGER, B. WOHLMUTH, AND T. DICKOPF, *Simultaneous reduced basis approximation of parameterized elliptic eigenvalue problems*, Esaim Math. Model. Numer. Anal., 51 (2017), pp. 443–465.
- [37] G. HU, H. XIE, AND F. XU, *A multilevel correction adaptive finite element method for Kohn–Sham equation*, J. Comput. Phys., 355 (2018), pp. 436–449.

- [38] S. JIA, H. XIE, M. XIE, AND F. XU, *A full multigrid method for nonlinear eigenvalue problems*, *Sci. China Math.*, 59 (2016), pp. 2037–2048.
- [39] P. KAZEMI AND M. ECKART, *Minimizing the gross-pitaevskii energy functional with the sobolev gradient—analytical and numerical results*, *International Journal of Computational Methods*, 7 (2010), pp. 453–475.
- [40] B. LANGWALLNER, C. ORTNER, AND E. SÜLI, *Existence and convergence results for the galerkin approximation of an electronic density functional*, *Math. Models Methods Appl. Sci.*, 20 (2010), pp. 2237–2265.
- [41] M. G. LARSON, *A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems*, *SIAM J. Numer. Anal.*, 38 (2000), pp. 608–625.
- [42] H. LI AND J. S. OVAL, *A posteriori eigenvalue error estimation for the schrödinger operator with the inverse square potential*, *Mathematics and Statistics Faculty Publications and Presentations*, (2015).
- [43] L. LIN, J. LU, AND L. YING, *Numerical methods for kohn–sham density functional theory*, *Acta Numerica*, 28 (2019), pp. 405–539.
- [44] Q. LIN AND H. XIE, *A multi-level correction scheme for eigenvalue problems*, *Math. Comput.*, 84 (2014), pp. 71–88.
- [45] X. LIU AND T. VEJCHODSKÝ, *Fully computable a posteriori error bounds for eigenfunctions*, *Numer. Math.*, (2022).
- [46] Y. MADAY AND G. TURINICI, *Error bars and quadratically convergent methods for the numerical simulation of the Hartree-Fock equations*, *Numer. Math.*, 94 (2003), pp. 739–770.
- [47] R. MCWEENY, *The density matrix in self-consistent field theory i. iterative construction of the density matrix*, *Proc. R. Soc. Lond. A Math. Phys. Sci.*, 235 (1956), pp. 496–509.
- [48] M. T. NAKAO, M. PLUM, AND Y. WATANABE, *Numerical Verification Methods and Computer-Assisted Proofs for Partial Differential Equations*, Springer, Singapore, 2019.
- [49] NEYMEYR, K. , *Solving mesh eigenproblems with multigrid efficiency*, *Methods for Scientific Computing. Variational problems and Applications*, (2003).
- [50] L. P. PITAEVSKII, S. STRINGARI, AND S. STRINGARI, *Bose-Einstein Condensation*, Clarendon Press, Apr. 2003.
- [51] M. R. RACHEVA AND A. B. ANDREEV, *Superconvergence postprocessing for eigenvalues*, *Comput. Methods Appl. Math.*, 2 (2002), pp. 171–185.
- [52] N. RAZA, S. SIAL, S. S. SIDDIQI, AND T. LOOKMAN, *Energy minimization related to the nonlinear schrödinger equation*, *J. Comput. Phys.*, 228 (2009), pp. 2572–2577.
- [53] C. C. J. ROOTHAAN, *New developments in molecular orbital theory*, *Rev. Mod. Phys.*, 23 (1951), pp. 69–89.

- [54] A. SCHMIDT, D. WITTHAR, AND B. HAASDONK, *Rigorous and effective a-posteriori error bounds for nonlinear problems—application to RB methods*, Adv. Comput. Math., 46 (2020), p. 32.
- [55] P. UPADHYAYA, E. JARLEBRING, AND E. H. RUBENSSON, *A density matrix approach to the convergence of the self-consistent field iteration*, Numer. Algebra Control Optim., 11 (2021), p. 99.
- [56] R. VERFURTH, *A Posteriori Error Estimation Techniques for Finite Element Methods (Numerical Mathematics and Scientific Computation)*, Oxford University Press, 1 ed., May 2013.
- [57] H. XIE AND M. XIE, *Computable error estimates for ground state solution of Bose–Einstein condensates*, J. Sci. Comput., 81 (2019), pp. 1072–1087.
- [58] F. XU AND Q. HUANG, *Cascadic adaptive finite element method for nonlinear eigenvalue problem based on complementary approach*, J. Comput. Appl. Math., 372 (2020), p. 112720.
- [59] J. XU AND A. ZHOU, *A two-grid discretization scheme for eigenvalue problems*, Math. Comput., 70 (1999), pp. 17–26.
- [60] F. ZHANG, F. SPIEGELMANN, E. SURAUD, V. FRAYSSÉ, R. POTEAU, R. GLOWINSKI, AND F. CHATELIN, *On the formation of transient (na19) 2 and (na20) 2 cluster dimers from molecular dynamics simulations*, Physics Letters A, 193 (1994), pp. 75–81.
- [61] F.-S. ZHANG, F. WANG, E. SURAUD, AND R. GLOWINSKI, *A distance dependent tight-binding molecular dynamics model to the collision and thermodynamical properties of nan*, Progress of Theoretical Physics Supplement, 138 (2000), pp. 72–77.
- [62] Z. ZHANG, *Exponential convergence of sobolev gradient descent for a class of nonlinear eigenproblems*, Commun. Math. Sci., 20 (2022), pp. 377–403.
- [63] A. ZHOU, *An analysis of finite-dimensional approximations for the ground state solution of Bose–Einstein condensates*, Nonlinearity, 17 (2003), p. 541.
- [64] ———, *Finite dimensional approximations for the electronic ground state solution of a molecular system*, Math. Methods Appl. Sci., 30 (2007), pp. 429–447.